Multidimensional generality of an integration model for personality impression formation

SUSUMU TAKAHASHI
Aichi University of Education, Kariya, Aichi, 448 Japan

A primary concern of the present study was a quantitative test of the dimensional generality of the weighted-average model of personality impression formation. Japanese college students were asked to make both likableness and intellectual judgments of stimulus persons described by two or four equal-valued personality adjectives. The two kinds of judgment both showed a strong set-size effect, with larger sets yielding more extreme responses. The results of the set-size analysis indicate that the weight parameters were constant across set sizes within each of the two response dimensions, thereby providing evidence to support the generality of the model on a multidimensional basis. The results also give some support to the cross-cultural generality of the weighted-average model.

Hamilton and Huffman (1971) tested the generality of the averaging model for personality impression formation on three response dimensions simultaneously: evaluation ("good/bad"), potency ("strong, domineering/weak, submissive"), and activity ("active/passive"). Their results showed that the averaging model holds for the dimensions of evaluation and potency, thereby confirming the model's generality on a multidimensional basis.

Although Hamilton and Huffman's study is one of the few that was concerned with the multidimensional generality of integration models, their test was only qualitative, designed to test between averaging and adding models. An adequate test of multidimensional generality requires an exact, quantitative test of the model. That is a primary concern of the present study.

Based on Rosenberg, Nelson, and Vivekananthan's (1968) distinction between social desirability and intellectual desirability, the present study tests the multidimensional generality of the weighted-average model on both likableness and intellectual judgments. Quantitatively, the weighted-average model can be tested by an analysis of the set-size effect, a well established phenomenon (Anderson, 1965, 1967; Takahashi, 1970, 1971), in which the greater the number of equal-valued information items, the more extreme the response.

At face value, the set-size effect seems to be consistent with an adding, not an averaging, model. However, Anderson (1965) indicated that the averaging model can account for the set-size effect in an exact, quantitative way, by assuming that the weight parameters are constant over set size. If the averaging model is true across response dimensions, the weight-constancy assumption should hold for both likableness and intellectual response dimensions. Since the set-size analysis is critical for the model, it needs to be conducted for both dimensions.

In addition to multidimensional generality, the present study will deal with the cross-cultural generality of the model. As summarized by Anderson (1974), most studies providing support for the averaging model have been conducted in America. There has been no research which gives support to the cross-cultural generality of the averaging model. Therefore, it seems worthwhile to assess model generality on a cross-cultural basis in the development of integration models. The important test of the cross-cultural generality of the set-size effect is that the weight parameters are constant in the Japanese as well as in the American culture.

METHOD

Preparation of Stimulus Materials

To collect normative data, 185 Japanese personality-adjective words were rated on the two dimensions of social desirability and intellectual desirability by 96 undergraduate students at Aichi University of Education. Both ratings were made on a 9-point scale with 9 labeled as highly desirable and 1 labeled as highly undesirable. The mean ratings were used to define the normative social and intellectual desirability scale values for each word. The correlation between the two sets of scale values was high ($r = .72$), replicating the results of Hamilton and Fallot (1974). While this does not mean that subjects cannot differentiate between the two content dimensions, it does place a restriction on selecting the stimulus words.

Based on the normative scale values, 185 words were divided into high (H), moderately high(M+), moderately low (M−), and low (L) subgroups of words for each of the two dimensions.

Table 1
Range of Scale Values of Trait Words in the H, M+, M−, and L Categories for Both Dimensions

Word Category	Intellectual Desirability		Social Desirability	
	Set Size 2	Set Size 4	Set Size 2	Set Size 4
H	6.86	6.85	7.32	7.28
M+	5.78	5.62	5.39	5.77
M−	4.32	4.22	3.74	3.65
L	3.02	3.10	2.26	2.25

Note–Trait words were rated on a 9-point scale, with 9 labeled as highly desirable and 1 labeled as highly undesirable.

By jointly considering the two sets of scale values, eight types of adjectives were distinguished. The four types of positive words were H_I-H_S (high intellectual-high social), H_I-M_S^+ (high intellectual-mildly high social), M_I^+-H_S (mildly high intellectual-high social), and M_I^+-M_S^+ (mildly high intellectual-mildly high social). The four types of negative words were M_I^--M_S^- (mildly low intellectual-mildly low social), M_I^--L_S (mildly low intellectual-low social), L_I-M_S^- (low intellectual-mildly low social), and L_I-L_S (low intellectual-low social). Examples of a typical word of each type are as follows: faithful (H_I-H_S), steady (H_I-M_S^+), cheerful (M_I^+-H_S), gentle (M_I^+-M_S^+), passive (M_I^--M_S^-), egoistic (L_I-M_S^-), insensible (M_I^--L_S), and random (L_I-L_S).

Because of high correlation between the two sets of scale values, only two levels of set size were used: two and four. Six words were chosen randomly from each of the eight types and were assigned to a set of two and a set of four adjectives, thereby forming a total of 16 sets. Repetition of words was not allowed. Table 1 lists the mean scale values for each dimension.

Procedure

The stimulus adjective sets were presented in a booklet consisting of a cover page and 16 pages, with each page containing a different set. Subjects were informed that the experiment was concerned with how people form a first impression of a strange person, and that in the booklet a number of persons were described by two or four personality-adjective words. The words describing each person were said to be attributed by acquaintances.

Subjects were asked to judge how much they thought they would like each person (likableness response dimension) and how intellectual they thought each person is (intellectual response dimension). Both ratings were made on a 21-point scale, with end points identified as 0 (highly unlikable or highly unintellectual), and 20 (highly likable or highly intellectual). It was emphasized that each rating was to be considered independently and the ratings were to be distributed across the entire 21-point scale. The two kinds of judgment were made in different response booklets. Half of the subjects were asked to make the likableness judgments before the intellectual judgments, and the other half were asked to rate in the reverse order. However, rating order had no effect in the statistical analyses and will be ignored below.

Subjects

Subjects were 30 male and 33 female undergraduate volunteers at Aichi University of Education in Japan who were different from those who served in prescaling of the normative scale values. They were tested three to five at a time.

RESULTS

The mean likableness and intellectual rating scores are listed in Tables 2 and 3, respectively. It can be seen

from these tables that the ratings become more extreme as set size increases for each judgment, and response variability is similar for the two types of judgment. There is a strong set-size effect for both response dimensions and for both positive and negative word sets.

The statistical analyses support the visual inspection. Results of an analysis of variance showed a significant effect of set size in the likableness ratings of positive word sets and negative word sets [$F(1,64) = 73.95$ and 88.96, respectively] and the intellectual ratings of positive word sets and negative word sets [$F(1,64) = 128.48$ and 31.09, respectively]. Of primary theoretical relevance is whether the weighted-average model can provide a quantitative account of the strong set-size effect. Accordingly, the two-dimensional impression ratings were subjected to the set-size analysis.

According to Anderson (1965), the weighted-average formulation can be written as:

Table 2
Mean Intellectual Rating Scores and Standard Deviations for Each Word Set

Word Set	Set Size 2		Set Size 4	
	Mean	SD	Mean	SD
Positive				
H_I-H_S	14.29	3.21	15.74	3.10
H_I-M_S^+	15.55	2.22	17.55	2.04
M_I^+-H_S	11.80	2.89	12.79	3.18
M_I^+-M_S^+	10.71	2.65	12.56	2.98
Negative				
M_I^--M_S^+	6.95	2.78	4.82	2.95
M_I^--L_S	8.64	3.24	7.27	3.28
L_I-M_S^-	4.18	3.06	3.85	2.62
L_I-L_S	3.91	2.94	3.02	2.30

Note–Intellectual rating was made on a 21-point scale with 0 labeled as highly unintellectual and 20 labeled as highly intellectual.

Table 3
Mean Likableness Rating Scores and Standard Deviations for Each Word Set

Word Set	Set Size 2		Set Size 4	
	Mean	SD	Mean	SD
Positive				
H-H	13.97	2.96	16.55	2.41
H-M+	11.79	3.24	11.92	3.53
M+-H	17.12	2.25	17.74	2.34
M+M+	11.29	2.87	13.58	2.99
Negative				
M−-M−	6.12	2.83	4.52	2.89
M−-L	2.79	2.53	1.33	1.59
L-M−	6.50	2.89	4.24	2.87
L-L	2.62	2.33	2.45	2.18

Note–Likableness rating was made on a 21-point scale with 0 labeled as highly unlikable and 20 labeled as highly likable.

$$Rk = \frac{kwA + (1 - w)Io}{kw + (1 - w)}.\qquad(1)$$

The numerator is the weighted sum of the scale values, A, of the k stimuli, and the initial or neutral impression, Io. The denominator is the sum of the weights and defines an averaging model. Quantitatively, the formulation assumes that w is constant across set sizes, even though the larger sets yield more extreme responses. Thus, the set-size analysis is a test of the constancy assumption of w over set sizes, and provides a powerful and critical test of the model. The weight values can be estimated from the observed responses by applying Equation 1 for each set size, where the values of A and Io are assigned on an a priori basis as the end points and the neutral or middle point of the response scale, respectively (Anderson, 1967). In this study, Io is set equal to 10, the nominal middle point of the scale for both response dimensions. A is set equal to the nominal end points, 20 for both H_I and H_S sets and 0 for both L_I and L_S sets.

For each subject, the numerical estimates of w were calculated for each set size of the intellectual ratings of H_I and L_I sets collapsed over social desirability and of the likableness ratings of H_S and L_S sets collapsed over intellectual desirability. M_I^+-M_S^+ and M_I^--M_S^- sets were omitted in this analysis because the scale values of mildly valued adjective sets cannot be assigned on an a priori basis. The mean w values are in Table 4, with subscripts used to denote set size.

Each of the four columns of Table 4 clearly shows that the w values are constant across set size within each response dimension, as the model requires. The statistical analysis confirms the visual inspection. None of the differences between w_2 and w_4 were significant. This replicates the findings of Anderson (1967). By using five levels of set size, Anderson found that constancy of the weight parameters holds over set sizes on the likableness response dimension. For likableness judgments with Set Sizes 2 and 4, the w values obtained here can be compared to those obtained by Anderson (shown in parentheses in Table 4). For positive sets, the values are exactly the same. For negative sets, the values obtained in the present study are higher than Anderson's values. Thus, in comparing these two particular studies, Japanese students place greater weight on negative traits than do American students. There is

Table 4
Mean w Estimates as a Function of Set Size and Set Value for Both Rating Dimensions

Estimated Quantity	Intellectual Judgment Set Value		Likableness Judgment Set Value	
	Positive	Negative	Positive	Negative
w_2	.40	.49	.45 (.45)	.66 (.43)
w_4	.45	.42	.46 (.46)	.64 (.47)

Note—Maximum value of w = 1.00. The numbers in parentheses are the weight estimates obtained by Anderson (1967, Table 1).

another small discrepancy between the results of the two studies. In Anderson's study, an end-effect response tendency was found, where subjects tended to use the scale end points for the largest sets, and consequently the largest w values were found for the largest sets. In the present study, however, this response tendency was not found. The difference may be due to the different set sizes included in each study.

While the small number of set sizes employed here may place some restriction on the generality of the present findings, the outcome of the set-size analysis indicates that the weighted-average model can account for the quantitative aspects of the present data for both response dimensions. This provides evidence to support the generality of the model on a multidimensional basis. Moreover, it can be said from a comparison of the present study and Anderson's (1967) study that the same weighted-averaging integration process holds for both Japanese and American subjects.

While the set-size analysis provides some evidence that the averaging model can account for the strong set-size effect found in both likableness and intellectual judgments, there still remains a doubt concerning to what extent subjects differentiate between social and intellectual desirability dimensions in making judgments of stimulus persons. If they could not differentiate between the two content dimensions, the parallel findings obtained here might merely reflect lack of differentiation.

For each response dimension, the ratings for the eight positive sets and the eight negative sets were separately subjected to an analysis of variance. The effect of social desirability on the intellectual judgments was significant only in the negative sets [$F(1,64) = 6.41$, $p < .05$]. On the other hand, the effect of intellectual desirability on the intellectual judgments was significant in both positive sets and negative sets [$F(1,64) = 255.06$ and 133.10, respectively]. The effect of intellectual desirability on the likableness judgments was significant only in the positive sets [$F(1,64) = 31.90$, $p < .001$], whereas the effect of social desirability on the likableness judgments was significant in both positive sets and negative sets [$F(1,64) = 168.26$ and 186.06, respectively]. While these results suggest in part some overlap between the two content dimensions, the effects were not systematic and were relatively small compared to the effect that the social desirability dimension had on likableness judgments and the effect that the intellectual dimension had on intellectual judgments. This replicates earlier findings (Friendly & Glucksberg, 1970; Hamilton & Fallot, 1974; Levin, 1973; Rosenberg et al., 1968).

DISCUSSION

The pattern of set-size effects was about the same for likableness and intellectual judgments. The statistical analyses show that the parallel findings were not due to lack of differentiation between social and intellectual

dimensions. Subjects give greater weight to the social desirability dimension in making likableness judgments and greater weight to the intellectual dimension when making intellectual judgments.

A primary concern of the present study was providing evidence that the weighted-average model can account for the set-size effect by showing that the w values are constant across set size within each of the two response dimensions, as the model requires. This basic assumption of the model is clearly favored in both likableness and intellectual judgments, thereby providing evidence to support the model's generality on a multidimensional basis.

Further generalization of the present data would require that similar results hold for other kinds of response dimensions, such as sociableness, cooperativeness, and kindness. However, whether the multidimensional generality of the averaging model holds for other types of stimuli is problematical. As far as an impression-formation process is concerned, the present findings might be generalizable to other types of stimuli, such as photographs and paragraphs describing behavioral characteristics. The present pattern of results might reappear in judgments on a more diverse array of stimulus materials than person descriptions.

The second interesting aspect of the present study concerns the cross-cultural generality of the averaging model. Some information on the cross-cultural generality of the model may be needed in order to develop the integration models more fully. There are two aspects of the cross-cultural generality of the set-size analysis. The most important is that the w values are constant across set sizes for each culture. The other aspect is associated with the size of the w values. Comparison of the present results with those obtained by Anderson (1967) relates to each of these two aspects. With respect to the former, the constancy assumption of the w values holds for both studies. This fact implies that the same weighted-averaging integration process holds in both Japan and America. The other aspect, the size of w values, is less relevant to the cross-cultural generality because the averaging model only requires the constancy of the weight parameters over set size and culture. While it is allowable to have different values for people of different cultures, the cross-cultural comparison of the size of the w values provides information on the extent to which people of different cultures depend on their own initial predispositions in making judgments of other persons. The weighted-average model for set-size effects includes the initial impression parameter, Io, representing the subject's

initial predisposition. The initial predisposition is reflected in the total impression, and the w values are determined postdictively from the observed total impression. Thus, if the initial predisposition differs among people of different cultures, this should be reflected in the observed response and should lead to different w values for the two cultures.

As noted earlier, the w estimates for positive traits were the same in the present study and in Anderson's (1967) study, but the w estimates for negative traits differed in the two studies. This suggests that there is no appreciable difference in the overall value and weight of initial predispositions of Japanese and American subjects. There may, however, be differences in the way Japanese and American subjects react to positive and negative traits in forming personality impressions.

REFERENCES

ANDERSON, N. H. Averaging versus aiding as a stimulus-combination rule in impression formation. *Journal of Experimental Psychology*, 1965, 70, 394-400.

ANDERSON, N. H. Averaging model analysis of set-size effect in impression formation. *Journal of Experimental Psychology*, 1967, 75, 158-165.

ANDERSON, N. H. *Basic experiments in person perception* (CHIP Report No. 44). San Diego: University of California, San Diego, Center for Human Information Processing, June 1974.

FRIENDLY, M. L., & GLUCKSBERG, S. On the description of subcultural lexicons: A multidimensional approach. *Journal of Personality and Social Psychology*, 1970, 14, 55-65.

HAMILTON, D. L., & FALLOT, R. D. Information salience as a weighting factor in impression formation. *Journal of Personality and Social Psychology*, 1974, 30, 444-448.

HAMILTON, D. L., & HUFFMAN, L. J. Generality of impression-formation processes for evaluative and nonevaluative judgments. *Journal of Personality and Social Psychology*, 1971, 20, 200-207.

LEVIN, I. P. Learning effects in information integration: Manipulation of cue validity in an impression formation task. *Memory & Cognition*, 1973, 1, 236-240.

ROSENBERG, S., NELSON, C., & VIVEKANANTHAN, P. S. A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, 1968, 9, 283-294.

TAKAHASHI, S. Analysis of weighted averaging model on integration of informations in personality impression formation. *Japanese Psychological Research*, 1970, 12, 154-162.

TAKAHASHI, S. Effect of the context upon personality impression formation. *Japanese Journal of Psychology*, 1971, 41, 307-313.